

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: SAMSON, MICHEL  
PARMENTIER, MARC  
VASSART, GILBERT  
LIBERT, FREDERICK

(ii) TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR  
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

(iii) NUMBER OF SEQUENCES: 17

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Knobbe, Martens, Olson & Bear  
(B) STREET: 620 Newport Center Drive 16th Floor  
(C) CITY: Newport Beach  
(D) STATE: CA  
(E) COUNTRY: U.S.A.  
(F) ZIP: 92660

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/833,752  
(B) FILING DATE: 9-APR-1997  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Altman, Daniel E  
(B) REGISTRATION NUMBER: 34,115  
(C) REFERENCE/DOCKET NUMBER:

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 792 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

(A) NAME/KEY: CDS  
 (B) LOCATION: 240..791

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAATTCCCCC AACAGAGCCA AGCTCTCCAT CTAGTGGACA GGGAGCTAG CAGCAAACCT	60
TCCCTCACT ACAAAACTTC ATTGCTTGGC CAAAAAGAGA GTTAATTCAA TGTAGACATC	120
TATGTAGGCA ATTAAAAACC TATTGATGTA TAAAACAGTT TGCATTCATG GAGGGCAACT	180
AAATACATTC TAGGACTTTA TAAAAGATCA CTTTTTATTT ATGCACAGGG TGGAACAAAG	239
ATG GAT TAT CAA GTG TCA AGT CCA ATC TAT GAC ATC AAT TAT TAT ACA	287
Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr	
1 5 10 15	
TCG GAG CCC TGC CAA AAA ATC AAT GTG AAG CAA ATC GCA GCC CGC CTC	335
Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu	
20 25 30	
CTG CCT CCG CTC TAC TCA CTG GTG TTC ATC TTT GGT TTT GTG GGC AAC	383
Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn	
35 40 45	
ATG CTG GTC ATC CTC ATC CTG ATA AAC TGC AAA AGG CTG AAG AGC ATG	431
Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met	
50 55 60	
ACT GAC ATC TAC CTG CTC AAC CTG GCC ATC TCT GAC CTG TTT TTC CTT	479
Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu	
65 70 75 80	
CTT ACT GTC CCC TTC TGG GCT CAC TAT GCT GCC GCC CAG TGG GAC TTT	527
Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Gln Trp Asp Phe	
85 90 95	
GGA ATAT ACA ATG TGT CAA CTC TTG ACA GGG CTC TAT TTT ATA GGC TTC	575
Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe	
100 105 110	
TTC TCT GGA ATC TTC TTC ATC ATC CTC CTG ACA ATC GAT AGG TAC CTG	623
Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu	
115 120 125	
GCT GTC GTC CAT GCT GTG TTT GCT TTA AAA GCC AGG ACG GTC ACC TTT	671
Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe	
130 135 140	
GGG GTG GTG ACA AGT GTG ATC ACT TGG GTG GTG GCT GTG TTT GCG TCT	719
Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser	
145 150 155 160	
CTC CCA GGA ATC ATC TTT ACC AGA TCT CAA AAA GAA GGT CTT CAT TAC	767
Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr	
165 170 175	
ACC TGC AGC TCT CAT TTT CCA TAC A	792
Thr Cys Ser Ser His Phe Pro Tyr	
180	

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1477 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 240..1295

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GAATTCCCCC AACAGAGCCA AGCTCTCCAT CTAGTGGACA GGGAGCTAG CAGCAAACCT	60
TCCCTTCACT ACAAAACTTC ATTGCTTGGC CAAAAAGAGA GTTAATTCAA TGTAGACATC	120
TATGTAGGCA ATTAAAAACC TATTGATGTA TAAAACAGTT TGCATTCATG GAGGGCAACT	180
AAATACATTC TAGGACTTTA TAAAAGATCA CTTTTTATTT ATGCACAGGG TGGAAACAAG	239
ATG GAT TAT CAA GTG TCA AGT CCA ATC TAT GAC ATC AAT TAT TAT ACA	287
Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr	
1 5 10 15	
TCG GAG CCC TGC CAA AAA ATC AAT GTG AAG CAA ATC GCA GCC CGC CTC	335
Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu	
20 25 30	
CTG CCT CCG CTC TAC TCA CTG GTG TTC ATC TTT GGT TTT GTG GGC AAC	383
Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn	
35 40 45	
ATG CTG GTC ATC CTC ATC CTG ATA AAC TGC AAA AGG CTG AAG AGC ATG	431
Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met	
50 55 60	
ACT GAC ATC TAC CTG CTC AAC CTG GCC ATC TCT GAC CTG TTT TTC CTT	479
Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu	
65 70 75 80	
CTT ACT GTC CCC TTC TGG GCT CAC TAT GCT GCC CAG TGG GAC TTT	527
Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe	
85 90 95	
GGA AAT ACA ATG TGT CAA CTC TTG ACA GGG CTC TAT TTT ATA GGC TTC	575
Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe	
100 105 110	
TTC TCT GGA ATC TTC ATC ATC CTC CTG ACA ATC GAT AGG TAC CTG	623
Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu	
115 120 125	
GCT GTC CAT GCT GTG TTT GCT TTA AAA GCC AGG ACG GTC ACC TTT	671
Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe	
130 135 140	
GGG GTG GTG ACA AGT GTG ATC ACT TGG GTG GTG GCT GTG TTT GCG TCT	719
Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser	
145 150 155 160	
CTC CCA GGA ATC ATC TTT ACC AGA TCT CAA AAA GAA GGT CTT CAT TAC	767
Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr	
165 170 175	

ACC TGC AGC TCT CAT TTT CCA TAC AGT CAG TAT CAA TTC TGG AAG AAT	815
Thr Cys Ser Ser His Phe Pro Tyr Ser Gln Tyr Gln Phe Trp Lys Asn	
180 185 190	
TTC CAG ACA TTA AAG ATA GTC ATC TTG GGG CTG GTC CTG CCG CTG CTT	863
Phe Gln Thr Leu Lys Ile Val Ile Leu Gly Leu Val Leu Pro Leu Leu	
195 200 205	
GTC ATG GTC ATC TGC TAC TCG GGA ATC CTA AAA ACT CTG CTT CGG TGT	911
Val Met Val Ile Cys Tyr Ser Gly Ile Leu Lys Thr Leu Leu Arg Cys	
210 215 220	
CGA AAT GAG AAG AAG AGG CAC AGG GCT GTG AGG CTT ATC TTC ACC ATC	959
Arg Asn Glu Lys Lys Arg His Arg Ala Val Arg Leu Ile Phe Thr Ile	
225 230 235 240	
ATG ATT GTT TAT TTT CTC TTC TGG GCT CCC TAC AAC ATT GTC CTT CTC	1007
Met Ile Val Tyr Phe Leu Phe Trp Ala Pro Tyr Asn Ile Val Leu Leu	
245 250 255	
CTG AAC ACC TTC CAG GAA TTC TTT GGC CTG AAT AAT TGC AGT AGC TCT	1055
Leu Asn Thr Phe Gln Glu Phe Phe Gly Leu Asn Asn Cys Ser Ser Ser	
260 265 270	
AAC AGG TTG GAC CAA GCT ATG CAG GTG ACA GAG ACT CTT GGG ATG ACG	1103
Asn Arg Leu Asp Gln Ala Met Gln Val Thr Glu Thr Leu Gly Met Thr	
275 280 285	
CAC TGC TGC ATC AAC CCC ATC ATC TAT GCC TTT GTC GGG GAG AAG TTC	1151
His Cys Cys Ile Asn Pro Ile Ile Tyr Ala Phe Val Gly Glu Lys Phe	
290 295 300	
AGA AAC TAC CTC TTA GTC TTC TTC CAA AAG CAC ATT GCC AAA CGC TTC	1199
Arg Asn Tyr Leu Leu Val Phe Phe Gln Lys His Ile Ala Lys Arg Phe	
305 310 315 320	
TGC AAA TGC TGT TCT ATT TTC CAG CAA GAG GCT CCC GAG CGA GCA AGC	1247
Cys Lys Cys Cys Ser Ile Phe Gln Gln Glu Ala Pro Glu Arg Ala Ser	
325 330 335	
TCA GTT TAC ACC CGA TCC ACT GGG GAG CAG GAA ATA TCT GTG GGC TTG	1295
Ser Val Tyr Thr Arg Ser Thr Gly Glu Gln Glu Ile Ser Val Gly Leu	
340 345 350	
TGACACGGAC TCAAGTGGGC TGGTGACCCA GTCAGAGTTG TGCACATGGC TTAGTTTCA	1355
TACACAGCCT GGGCTGGGG TNGGTTGGNN GAGGTCTTT TTAAAAGGAA GTTACTGTTA	1415
TAGAGGGTCT AAGATTTCATC CATTATTTG GCATCTGTT AAAGTAGATT AGATCCGAAT	1475
TC	1477

## (2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1442 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
 

- (A) NAME/KEY: CDS

(B) LOCATION: 240..884

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAATTCCCC AACAGAGCCA AGCTCTCCAT	CTAGTGGACA GGGAGCTAG CAGCAAACCT	60
TCCCTTCACT ACAAAACTTC ATTGCTTGGC	CAAAAAGAGA GTTAATTCAA TGTAGACATC	120
TATGTAGGCA ATTAAAAACC TATTGATGTA	AAAACAGTT TGCATTCATG GAGGGCAACT	180
AAATACATTC TAGGACTTTA TAAAAGATCA	CTTTTATTT ATGCACAGGG TGGAACAAG	239
ATG GAT TAT CAA GTG TCA AGT CCA ATC TAT GAC ATC AAT TAT TAT ACA		287
Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr		
1 5 10 15		
TCG GAG CCC TGC CAA AAA ATC AAT GTG AAG CAA ATC GCA GCC CGC CTC		335
Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu		
20 25 30		
CTG CCT CCG CTC TAC TCA CTG GTG TTC ATC TTT GGT TTT GTG GGC AAC		383
Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn		
35 40 45		
ATG CTG GTC ATC CTC ATC CTG ATA AAC TGC AAA AGG CTG AAG AGC ATG		431
Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met		
50 55 60		
ACT GAC ATC TAC CTG CTC AAC CTG GCC ATC TCT GAC CTG TTT TTC CTT		479
Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu		
65 70 75 80		
CTT ACT GTC CCC TTC TGG GCT CAC TAT GCT GCC CAG TGG GAC TTT		527
Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe		
85 90 95		
GGA AAT ACA ATG TGT CAA CTC TTG ACA GGG CTC TAT TTT ATA GGC TTC		575
Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe		
100 105 110		
TTC TCT GGA ATC TTC TTC ATC ATC CTC CTG ACA ATC GAT AGG TAC CTG		623
Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu		
115 120 125		
GCT GTC CAT GCT GTG TTT GCT TTA AAA GCC AGG ACG GTC ACC TTT		671
Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe		
130 135 140		
GGG GTG GTG ACA AGT GTG ATC ACT TGG GTG GTG GCT GTG TTT GCG TCT		719
Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser		
145 150 155 160		
CTC CCA GGA ATC ATC TTT ACC AGA TCT CAA AAA GAA GGT CTT CAT TAC		767
Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr		
165 170 175		
ACC TGC AGC TCT CAT TTT CCA TAC ATT AAA GAT AGT CAT CTT GGG GCT		815
Thr Cys Ser Ser His Phe Pro Tyr Ile Lys Asp Ser His Leu Gly Ala		
180 185 190		
GGT CCT GCC GCT GCT TGT CAT GGT CAT CTG CTA CTC GGG AAT CCT AAA		863
Gly Pro Ala Ala Ala Cys His Gly His Leu Leu Leu Gly Asn Pro Lys		
195 200 205		
AAC TCT GCT TCG GTG TCG AAA TGAGAAGAG AGGCACAGGG CTGTGAGGCT		914
Asn Ser Ala Ser Val Ser Lys		
210 215		
TATCTTCACC ATCATGATTG TTTATTTCT CTTCTGGCT CCCTACAACA TTGTCCTTCT		974

CCTGAACACC	TTCCAGGAAT	TCTTGCCCT	GAATAATTGC	AGTAGCTCTA	ACAGGTTGGA	1034
CCAAGCTATG	CAGGTGACAG	AGACTCTTGG	GATGACGCAC	TGCTGCATCA	ACCCCACAT	1094
CTATGCCCTT	GTCGGGGAGA	AGTTCAAGAAA	CTACCTCTTA	GTCTTCTTCC	AAAAGCACAT	1154
TGCCAAACGC	TTCTGCAAAT	GCTGTTCTAT	TTTCCAGCAA	GAGGCTCCCG	AGCGAGCAAG	1214
CTCAGTTAC	ACCCGATCCA	CTGGGGAGCA	GGAAATATCT	GTGGGCTTGT	GACACGGACT	1274
CAAGTGGGCT	GGTGACCCAG	TCAGAGTTGT	GCACATGGCT	TAGTTTCTAT	ACACAGCCTG	1334
GGCTGGGGGT	GGTTGGGAGG	TCTTTTTAA	AAGGAAGTTA	CTGTTATAGA	GGGTCTAAGA	1394
TTCATCCATT	TATTTGGCAT	CTGTTAAAG	TAGATTAGAT	CCGAATTC		1442

## (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 184 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

iii) MOLECULE TYPE: protein

xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met	Asp	Tyr	Gln	Val	Ser	Ser	Pro	Ile	Tyr	Asp	Ile	Asn	Tyr	Tyr	Thr
1				5					10					15	
Ser	Glu	Pro	Cys	Gln	Lys	Ile	Asn	Val	Lys	Gln	Ile	Ala	Ala	Arg	Leu
2					20				25					30	
Leu	Pro	Pro	Leu	Tyr	Ser	Leu	Val	Phe	Ile	Phe	Gly	Phe	Val	Gly	Asn
3						35			40					45	
Met	Leu	Val	Ile	Leu	Ile	Leu	Ile	Asn	Cys	Lys	Arg	Leu	Lys	Ser	Met
4						50			55					60	
Thr	Asp	Ile	Tyr	Leu	Leu	Asn	Leu	Ala	Ile	Ser	Asp	Leu	Phe	Phe	Leu
5						65			70					80	
Leu	Thr	Val	Pro	Phe	Trp	Ala	His	Tyr	Ala	Ala	Ala	Gln	Trp	Asp	Phe
6						85				90				95	
Gly	Asn	Thr	Met	Cys	Gln	Leu	Leu	Thr	Gly	Leu	Tyr	Phe	Ile	Gly	Phe
7						100			105					110	
Phe	Ser	Gly	Ile	Phe	Phe	Ile	Ile	Leu	Leu	Thr	Ile	Asp	Arg	Tyr	Leu
8						115			120					125	
Ala	Val	Val	His	Ala	Val	Phe	Ala	Leu	Lys	Ala	Arg	Thr	Val	Thr	Phe
9						130			135					140	
Gly	Val	Val	Thr	Ser	Val	Ile	Thr	Trp	Val	Val	Ala	Val	Phe	Ala	Ser
10						145			150					160	
Leu	Pro	Gly	Ile	Ile	Phe	Thr	Arg	Ser	Gln	Lys	Glu	Gly	Leu	His	Tyr
11						165			170					175	
Thr	Cys	Ser	Ser	His	Phe	Pro	Tyr								
12						180									

## (2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 352 amino acids
  - (B) TYPE: amino acid

## (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr  
 1 5 10 15  
 Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu  
 20 25 30  
 Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn  
 35 40 45  
 Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met  
 50 55 60  
 Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu  
 65 70 75 80  
 Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Gln Trp Asp Phe  
 85 90 95  
 Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe  
 100 105 110  
 Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu  
 115 120 125  
 Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe  
 130 135 140  
 Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser  
 145 150 155 160  
 Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr  
 165 170 175  
 Thr Cys Ser Ser His Phe Pro Tyr Ser Gln Tyr Gln Phe Trp Lys Asn  
 180 185 190  
 Phe Gln Thr Leu Lys Ile Val Ile Leu Gly Leu Val Leu Pro Leu Leu  
 195 200 205  
 Val Met Val Ile Cys Tyr Ser Gly Ile Leu Lys Thr Leu Leu Arg Cys  
 210 215 220  
 Arg Asn Glu Lys Lys Arg His Arg Ala Val Arg Leu Ile Phe Thr Ile  
 225 230 235 240  
 Met Ile Val Tyr Phe Leu Phe Trp Ala Pro Tyr Asn Ile Val Leu Leu  
 245 250 255  
 Leu Asn Thr Phe Gln Glu Phe Phe Gly Leu Asn Asn Cys Ser Ser Ser  
 260 265 270  
 Asn Arg Leu Asp Gln Ala Met Gln Val Thr Glu Thr Leu Gly Met Thr  
 275 280 285  
 His Cys Cys Ile Asn Pro Ile Ile Tyr Ala Phe Val Gly Glu Lys Phe  
 290 295 300  
 Arg Asn Tyr Leu Leu Val Phe Phe Gln Lys His Ile Ala Lys Arg Phe  
 305 310 315 320  
 Cys Lys Cys Cys Ser Ile Phe Gln Gln Glu Ala Pro Glu Arg Ala Ser  
 325 330 335  
 Ser Val Tyr Thr Arg Ser Thr Gly Glu Gln Glu Ile Ser Val Gly Leu  
 340 345 350

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 215 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr  
 1 5 10 15  
 Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu  
 20 25 30  
 Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn  
 35 40 45  
 Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met  
 50 55 60  
 Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu  
 65 70 75 80  
 Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Gln Trp Asp Phe  
 85 90 95  
 Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe  
 100 105 110  
 Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu  
 115 120 125  
 Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe  
 130 135 140  
 Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser  
 145 150 155 160  
 Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr  
 165 170 175  
 Thr Cys Ser Ser His Phe Pro Tyr Ile Lys Asp Ser His Leu Gly Ala  
 180 185 190  
 Gly Pro Ala Ala Ala Cys His Gly His Leu Leu Leu Gly Asn Pro Lys  
 195 200 205  
 Asn Ser Ala Ser Val Ser Lys  
 210 215

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: None

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Leu Ser Thr Ser Arg Ser Arg Phe Ile Arg Asn Thr Asn Glu Ser  
 1 5 10 15  
 Gly Glu Glu Val Thr Thr Phe Phe Asp Tyr Asp Tyr Gly Ala Pro Cys  
 20 25 30  
 His Lys Phe Asp Val Lys Gln Ile Gly Ala Gln Leu Leu Pro Pro Leu  
 35 40 45  
 Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn Met Leu Val Val  
 50 55 60  
 Leu Ile Leu Ile Asn Cys Lys Leu Lys Cys Leu Thr Asp Ile Tyr  
 65 70 75 80  
 Leu Leu Asn Leu Ala Ile Ser Asp Leu Leu Phe Ile Ile Thr Leu Pro  
 85 90 95  
 Leu Trp Ala His Ser Ala Ala Asn Glu Trp Val Phe Gly Asn Ala Met  
 100 105 110  
 Cys Lys Leu Phe Thr Gly Leu Tyr His Ile Gly Tyr Phe Gly Gly Ile  
 115 120 125  
 Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu Ala Ile Val His  
 130 135 140  
 Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe Gly Val Val Thr  
 145 150 155 160  
 Ser Val Ile Thr Trp Leu Val Ala Val Phe Ala Ser Val Pro Gly Ile  
 165 170 175  
 Ile Phe Thr Lys Cys Gln Lys Glu Asp Ser Val Tyr Val Cys Gly Pro  
 180 185 190  
 Tyr Phe Pro Arg Gly Trp Asn Asn Phe His Thr Ile Met Arg Asn Ile  
 195 200 205  
 Leu Gly Leu Val Leu Pro Leu Leu Ile Met Val Ile Cys Tyr Ser Gly  
 210 215 220  
 Ile Leu Lys Thr Leu Leu Arg Cys Arg Asn Glu Lys Lys Arg His Arg  
 225 230 235 240  
 Ala Val Arg Val Ile Phe Thr Ile Met Ile Val Tyr Phe Leu Phe Trp  
 245 250 255  
 Thr Pro Tyr Asn Ile Val Ile Leu Leu Asn Thr Phe Gln Glu Phe Phe  
 260 265 270  
 Gly Leu Ser Asn Cys Glu Ser Thr Ser Gln Leu Asp Gln Ala Ile Gln  
 275 280 285  
 Val Thr Glu Thr Leu Gly Met Thr His Cys Cys Ile Asn Pro Ile Ile  
 290 295 300  
 Tyr Ala Phe Val Gly Glu Lys Phe Arg Arg Tyr Ile Ser Val Phe Phe  
 305 310 315 320  
 Arg Lys His Ile Xaa Xaa Xaa Phe Cys Lys Gln Cys Pro Val Phe Tyr  
 325 330 335  
 Arg Glu Thr Val Asp Gly Val Thr Ser Thr Asn Thr Pro Ser Thr Gly  
 340 345 350  
 Glu Gln Glu Val Ser Ala Gly Leu  
 355 360

## (2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 355 amino acids
- (B) TYPE: amino acid

(C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Thr Thr Ser Ile Asp Thr Val Glu Thr Phe Gly Thr Thr Ser Tyr  
 1 5 10 15  
 Tyr Asp Asp Val Gly Leu Leu Cys Glu Lys Ala Asp Thr Arg Ala Leu  
 20 25 30  
 Met Ala Gln Phe Val Pro Pro Leu Tyr Ser Leu Val Phe Thr Val Gly  
 35 40 45  
 Leu Ile Gly Asn Val Val Val Val Met Ile Leu Ile Lys Tyr Arg Arg  
 50 55 60  
 Ile Arg Ile Met Thr Asn Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp  
 65 70 75 80  
 Leu Leu Phe Ile Val Thr Leu Pro Phe Trp Thr His Tyr Val Arg Gly  
 85 90 95  
 His Asn Trp Val Phe Gly His Gly Met Cys Asn Leu Ile Ser Gly Phe  
 100 105 110  
 Tyr His Thr Gly Leu Tyr Ser Glu Ile Phe Phe Ile Ile Leu Leu Thr  
 115 120 125  
 Ile Asp Arg Tyr Leu Ala Ile Val His Ala Val Phe Ala Ile Arg Ala  
 130 135 140  
 Arg Thr Val Thr Phe Gly Val Ile Thr Ser Ile Val Thr Trp Gly Ile  
 145 150 155 160  
 Ala Val Ile Ala Ala Leu Pro Glu Phe Ile Phe Tyr Glu Thr Glu Glu  
 165 170 175  
 Leu Phe Glu Glu Thr Ile Cys Ser Ala Leu Tyr Pro Glu Asp Thr Val  
 180 185 190  
 Tyr Ser Trp Arg His Phe His Thr Ile Arg Met Thr Ile Phe Cys Leu  
 195 200 205  
 Val Leu Pro Leu Leu Val Met Ala Ile Cys Tyr Thr Gly Ile Ile Lys  
 210 215 220  
 Thr Leu Leu Arg Cys Pro Xaa Xaa Xaa Lys Tyr Lys Ala Ile Arg Leu  
 225 230 235 240  
 Ile Phe Val Ile Met Ala Val Phe Phe Ile Glu Trp Thr Pro Tyr Asn  
 245 250 255  
 Val Ala Ile Leu Ile Ser Ser Tyr Gln Ser Leu Leu Phe Gly Asn Asn  
 260 265 270  
 Cys Glu Arg Ser Lys His Leu Asp Leu Val Met Ile Val Thr Glu Val  
 275 280 285  
 Ile Ala Tyr Ser His Cys Cys Met Asn Glu Val Ile Tyr Ala Phe Val  
 290 295 300  
 Gly Glu Arg Phe Arg Lys Tyr Ile Arg His Phe Phe His Arg His Leu  
 305 310 315 320  
 Leu Met His Leu Gly Arg Tyr Ile Pro Phe Leu Pro Xaa Xaa Xaa Ile  
 325 330 335  
 Glu Arg Ile Ser Ser Val Ser Pro Ser Thr Ala Glu Pro Glu Ile Ser  
 340 345 350  
 Ile Val Phe

355

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 355 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met	Glu	Thr	Pro	Asn	Thr	Thr	Glu	Asp	Tyr	Asp	Thr	Thr	Glu	Phe	
1	5						10						15		
Asp	Tyr	Gly	Asp	Ala	Thr	Pro	Cys	Gln	Lys	Val	Asn	Glu	Arg	Ala	Phe
20							25						30		
Gly	Ala	Gln	Leu	Leu	Pro	Pro	Leu	Tyr	Ser	Leu	Val	Phe	Val	Ile	Gly
35							40						45		
Leu	Val	Gly	Asn	Ile	Leu	Val	Val	Leu	Val	Leu	Val	Gln	Tyr	Lys	Arg
50							55						60		
Leu	Lys	Asn	Met	Thr	Ser	Ile	Tyr	Leu	Leu	Asn	Leu	Ala	Ile	Ser	Asp
65							70						75		80
Leu	Leu	Phe	Ile	Phe	Thr	Leu	Pro	Phe	Trp	Ile	Asp	Tyr	Lys	Leu	Lys
							85						90		95
Asp	Asp	Trp	Val	Phe	Gly	Asp	Ala	Met	Cys	Lys	Ile	Ile	Ser	Gly	Phe
100							105						110		
Tyr	Tyr	Thr	Gly	Leu	Tyr	Ser	Glu	Ile	Phe	Phe	Ile	Ile	Leu	Leu	Thr
115							120						125		
Ile	Asp	Arg	Tyr	Leu	Ala	Ile	Val	His	Ala	Val	Phe	Ala	Ile	Arg	Ala
130							135						140		
Arg	Thr	Val	Thr	Phe	Gly	Val	Ile	Thr	Ser	Ile	Ile	Trp	Ala	Ile	
145							150						155		160
Ala	Ile	Ile	Ala	Ser	Met	Pro	Gly	Leu	Tyr	Phe	Ser	Lys	Thr	Gln	Trp
							165						170		175
Glu	Phe	Thr	His	His	Thr	Cys	Ser	Leu	His	Phe	Pro	His	Glu	Ser	Leu
							180						185		190
Arg	Glu	Trp	Lys	Leu	Phe	Gln	Ala	Leu	Lys	Leu	Asn	Leu	Phe	Gly	Leu
195							200						205		
Val	Leu	Pro	Leu	Leu	Val	Met	Ile	Ile	Cys	Tyr	Ile	Gly	Ile	Ile	Lys
210							215						220		
Ile	Leu	Leu	Arg	Arg	Pro	Asn	Glu	Lys	Lys	Ser	Lys	Ala	Val	Arg	Leu
225							230						235		240
Ile	Phe	Val	Ile	Met	Ile	Ile	Phe	Phe	Leu	Phe	Trp	Ile	Pro	Tyr	Asn
							245						250		255
Leu	Thr	Ile	Ile	Ile	Ser	Val	Phe	Gln	Asp	Phe	Leu	Phe	Thr	His	Glu
							260						265		270
Cys	Glu	Gln	Ser	Arg	His	Leu	Asp	Leu	Ala	Val	Gln	Val	Thr	Glu	Val
275							280						285		
Ile	Ala	Tyr	Thr	His	Cys	Cys	Val	Asn	Glu	Val	Ile	Tyr	Ala	Phe	Val
290							295						300		

Gly Glu Arg Phe Arg Lys Tyr Ile Arg Gln Leu Glu His Arg Arg Val  
 305 310 315 320  
 Ala Val His Leu Val Lys Trp Leu Pro Phe Leu Ser Val Asp Arg Ile  
 325 330 335  
 Glu Arg Val Ser Ser Thr Ser Pro Ser Thr Gly Glu His Glu Ile Ser  
 340 345 350  
 Ala Gly Phe  
 355

## (2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Asn Pro Thr Asp Ile Ala Asp Thr Thr Leu Asp Glu Ser Ile Tyr  
 1 5 10 15  
 Ser Asn Tyr Tyr Leu Tyr Glu Ser Ile Pro Lys Pro Cys Thr Lys Glu  
 20 25 30  
 Gly Ile Lys Ala Phe Gly Glu Leu Phe Leu Pro Pro Leu Tyr Ser Leu  
 35 40 45  
 Val Glu Val Phe Gly Leu Ile Gly Asn Ser Val Val Val Leu Val Leu  
 50 55 60  
 Phe Lys Tyr Lys Arg Ile Arg Ser Met Thr Asp Val Tyr Leu Leu Asn  
 65 70 75 80  
 Leu Ala Ile Ser Asp Leu Leu Phe Val Phe Ser Leu Pro Phe Trp Gly  
 85 90 95  
 Tyr Tyr Ala Ala Asp Gln Trp Val Phe Gly Leu Gly Ile Cys Lys Met  
 100 105 110  
 Ile Ser Trp Met Tyr Leu Val Gly Phe Tyr Ser Gly Ile Phe Phe Val  
 115 120 125  
 Met Ile Met Ser Ile Asp Arg Tyr Leu Ala Ile Val His Ala Val Glu  
 130 135 140  
 Xaa Xaa Xaa Ala Arg Thr Ile Ile Tyr Gly Val Ile Thr Ser Leu Ala  
 145 150 155 160  
 Thr Trp Ser Val Ala Val Phe Ala Ser Leu Pro Gly Phe Ile Phe Ser  
 165 170 175  
 Thr Cys Tyr Thr Glu Arg Asn His Thr Tyr Cys Lys Thr Lys Tyr Ser  
 180 185 190  
 Leu Asn Ser Thr Thr Trp Lys Val Leu Ser Ser Leu Glu Ile Asn Ile  
 195 200 205  
 Leu Gly Leu Val Ile Pro Leu Gly Ile Met Leu Phe Cys Tyr Ser Met  
 210 215 220  
 Ile Ile Arg Thr Leu Gln His Cys Lys Asn Glu Lys Lys Asn Lys Ala  
 225 230 235 240  
 Val Lys Met Ile Phe Ala Val Val Leu Phe Leu Gly Phe Trp Thr

245	250	255	
Pro Tyr Asn Ile Val Leu Phe Leu Glu Thr Leu Val Glu Leu Glu Val			
260	265	270	
Ile Gln Asp Cys Thr Phe Glu Arg Tyr Leu Asp Tyr Ala Ile Gln Ala			
275	280	285	
Thr Glu Thr Leu Ala Phe Val His Cys Cys Leu Asn Pro Ile Ile Tyr			
290	295	300	
Phe Phe Leu Gly Glu Lys Phe Arg Lys Tyr Ile Ile Gln Leu Phe Lys			
305	310	315	320
Xaa Xaa Xaa Gly Leu Phe Val Ile Cys Gln Tyr Cys Gly Leu Leu Gln			
325	330	335	
Ile Tyr Ser Ala Asp Thr Pro Ser Ser Tyr Thr Gln Ser Thr Met			
340	345	350	
Asp His Asp Leu His Asp Ala Leu			
355	360		

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

The Cys Ser Ser His Phe Pro Tyr Ser Gln Tyr Gln Phe Trp Lys Asn			
14	5	10	15
Pho Gln Thr Leu Lys Ile Val Ile Leu Gly Leu Val Leu Pro Leu Leu			
20	25	30	
Val Met Val Ile Cys Tyr Ser Gly Ile Leu Lys Thr Leu Leu Arg Cys			
35	40	45	
Arg Asn Glu Lys Lys Arg			
50			

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 147 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTTCCATACA GTCAGTATCA ATTCTGGAAG AATTCCAGA CATTAAGAT AGTCATCTTG	60
GGGCTGGTCC TGCCGCTGCT TGTCATGGTC ATCTGCTACT CGGGAATCCT AAAAACTCTG	120
CTTCGGTGTC GAAATGAGAA GAAGAGG	147

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Phe Pro Tyr Ile Lys Asp Ser His Leu Gly Ala Gly Pro Ala Ala Ala  
 1 5 10 15  
 Cys His Gly His Leu Leu Leu Gly Asn Pro Lys Asn Ser Ala Ser Val  
 20 25 30  
 Ser Lys

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TCAGGATCC AAGATGGATT ATCAAGT

27

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTGATCTAGA GCCATGTGCA CAACTCT

27

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCTGGCTGTC GTCCATGCTG

20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTGATCTAGA GCCATGTGCA CAACTCT

27

On the 10th inst. the author and crew took the boat and went to the mouth of the river, and were there for a short time.